A FOLK THEOREM FOR COMPETING MECHANISMS

MICHAEL PETERS
UNIVERSITY OF BRITISH COLUMBIA
AND CRISTIAN TRONCOSO VALVERDE
UNIVERSIDAD DIEGO PORTALES MANUEL RODRÍGUEZ

Abstract. We provide a partial characterization of the set of outcome functions that can be supported as Bayesian equilibrium in the recommendation game described in Yamashita (Econometrica 2010). Despite the fact that communication is private in Yamashita's game, set of outcome functions that can be supported is effectively as large as the set supportable by a mechanism designer. In particular, we show how to support random and correlated outcomes, and illustrate how to ensure that the information used by different principals is consistent.

Many outcome functions can typically be supported as equilibria in competing mechanism games. Some of these outcomes look quite 'collusive'. The reason for this is that competing mechanism games often provide players the opportunity to make what they do conditional on what other players do. This allows players to support collusive outcomes by writing contracts that commit them to react whenever an opponent deviates from a putative equilibrium outcome. A complete characterization of supportable outcomes is provided in Peters (2010). Part of this characterization is the description of a reciprocal contracting game that can be used to support any (feasible) outcome.

Since the set of supportable outcome functions is large, it is natural to search for alternative extensive form contracting games in which players ability to commit or to communicate is restricted. One such contracting game was proposed by Yamashita (2010), The logic of his game is straightforward. Each principal commits to a mechanism that simply asks agents what it should do. If the majority of the agents' recommendations agree, the principal commits himself to carry out the recommendation.

Then, to support a particular outcome function as an equilibrium, principals offer recommendation mechanisms, and on the equilibrium path, agents unanimously recommend that each principal carry out

April 21, 2012.

his part of the agreement. Should any principal deviate and try to offer something other than a recommendation mechanism, the agents unanimously recommend that the others punish the deviator. The reason the agents are willing to do this is because they expect all the other agents to do it, and believe they will be ignored if they don't do likewise.

One reason that Yamashita's recommendation game is interesting is that players communicate their type reports privately. It is interesting to ask whether this restriction might limit the set of outcomes that would otherwise be supported as equilibrium. What is perhaps surprising about our result, is that private communication imposes essentially no restrictions on the set of supportable outcomes.

There are a number of reasons we need to write a paper on this instead of referring to Yamashita. First of all, though he explains perfectly well how competing mechanisms can be used to support multiple outcomes, he doesn't provide an explicit theorem characterizing the things that are supportable. Characterization isn't really the point of his paper. When he describes what a characterization might look like, he describes a 'value' that imposes a lower bound on principals' payoffs from supportable outcomes. This 'value' is the lowest payoff that the principal attains from any mechanism he can offer in any continuation equilibrium against any array of mechanisms of the other players. Since the calculation of these values basically requires the calculation of all equilibrium strategy rules, the 'characterization' is really nothing more than a restatement of the definition of equilibrium. So our primary objective in this paper is to turn this argument into a representation that can be used to compare his result to the rest of the literature.

One of the difficulties that arise in doing this is that Yamashita restricts players to pure strategies and non-random mechanisms. This is sensible for expositional reasons in his paper, but here we want to illustrate formally how to handle randomization. One benefit of our approach is that it shows how principals can use recommendation mechanisms to implement correlated actions.

The second difficulty has to do with private communication. What agents 'recommend' to principals in Yamashita's game is a direct mechanism. In the course of the operation of this direct mechanism a principal communicates privately with each of this agents, which determines the principal's own action. We show how to tie these private communications together in such a way that principals can coordinate their action choices.

Finally, Yamashita limits commitment ability to a group of uninformed principals who deal with informed agents who have no commitment power at all, and who make no direct choices beyond the messages that they send to principals. We show how to extend his approach to problems with informed principals and to situations in which all participants have commitment power.

Our main result is a characterization of a set of outcome functions that are supportable in his game. For complete information games, our results show how the folk theorem like results in Kalai, Kalai, Lehrer, and Samet (2010) can be extended to arbitrary numbers of players. For games of incomplete information, we show that this set is effectively as large as the set of outcome functions supportable by a centralized mechanism designer. We explain what we mean by 'effectively' below.

After presenting our formalism and results in the next few sections, we return to discuss some of the shortcomings of the recommendation game. In particular, we discuss the use of Bayesian equilibrium as a solution concept, and explain why we cannot give a full characterization of the set of outcome functions that can be supported as equilibrium in this game.

1. Fundamentals

There are $n \geq 4$ players. We sometimes write N to represent the set of players. Player i must choose an action a_i from a finite set A_i. Let $a = \{a_1, \ldots, a_n\}$ be an array of actions in $A = A_1 \times \cdots \times A_n$. $A_{-i} = \prod_{j \neq i} A_j$.

Each player i has a privately observed payoff type θ_i drawn from a finite set Θ. Payoffs are given by $u_i : A \times \Theta^n \to \mathbb{R}$. Players have expected utility preferences over actions.

Let P_i, P_{-i}, and P be the set of probability distributions on A_i, A_{-i}, and A respectively. A typical element $p \in P$ is a vector with p_k equal to the probability that the k^{th} element in A occurs, where the set A is indexed in some arbitrary fashion.

Let $q : \Theta^n \to P$ be an allocation rule. In what follows we slightly abuse notation by writing $u_i(q, \theta)$ instead of $\sum_{a \in A} q_a u_i(a, \theta)$. We are interested in allocation rules that are incentive compatible and individually rational. Incentive compatibility means

$$\mathbb{E}\left\{u_i\left(q\left(\theta\right), \theta\right) | \theta_i\right\} \geq \mathbb{E}\left\{u_i\left(q\left(\theta_i', \theta_{-i}\right), \theta\right) | \theta_i\right\} \tag{1.1}$$

for each $i \in N$, and $\theta_i' \in \Theta_i$. Individual rationality means that for each player i there is a punishment $p^i : \Theta_{-i} \to P_{-i}$ such that for every θ_i

$$\mathbb{E}\left\{u_i\left(q\left(\theta_i, \theta_{-i}\right), \left(\theta_i, \theta_{-i}\right)\right) | \theta_i\right\} \geq$$

$$(1.2) \qquad \max_{a_i} E\left\{u_i\left(a_i, p^i\left(\theta_{-i}\right), \left(\theta_i, \theta_{-i}\right)\right) | \theta_i\right\}.$$

With complete information, an allocation is individually rational if and only if it provides each player with an expected payoff that exceeds his or her *minmax value,* defined for player i as

$$(1.3) \qquad u_i^* \equiv \min_{p_{-i} \in P_{-i}} \max_{a_i \in A_i} u_i\left(a_i, p^i\right).$$

Again, with complete information the punishment

$$p_{-i}^* \in \arg \min_{p_{-i} \in P_{-i}} \max_{a_i} u_i\left(a_i, p^i\right)$$

can be used to support all implementable allocations.

Notice that when constructing a punishment, or a minmax value, punishers are allowed to correlate their punishments. This is appropriate for a mechanism designer who can enforce contracts and correlate actions among agents who have agreed to participate.

2. Recommendation Game

One of the things that makes competing mechanism games challenging is specifying exactly what message spaces and mechanisms are feasible for players. Since our objective here is to study the implications of private communication, we use a very narrow interpretation of what the set of feasible mechanisms is.

Let Γ_i be the set of all measurable mappings from $(\Theta \times [0,1])^{n(n-1)} \to A_i$. We are going to let the message space $\mathcal{M}_i$ for player i be $\Gamma_i \times (\Theta \times [0,1])^n$. The set of mechanisms for player i is then going to be the set of all measurable mappings $\mathcal{R}_i$ from $(\mathcal{M}_i)^{n-1}$ into A_i. The game then takes place in two stages. In the first, players publicly announce their mechanisms, in the second players privately send messages.

Let us explain. We are going to have each player i offer a mechanism that asks the other players to make a recommendation and report types and correlating messages. Up to some qualifications that we specify below, we describe equilibria in which each principal commits himself to follow the other players' recommendations provided they all agree. Recommendations are mappings that look pretty much like direct mechanisms, save for the additional correlating messages.

The complication we need to address is to ensure that the type report (and correlating message) that another player provides is the same type report that player sent to every third player. The way we are going to deal with this is to have players send messages in $(\Theta \times [0,1])^n$ over two rounds. Each player will report his own type and correlating message in the first round, then report the messages that he heard from the other types on the second round.

As the process of reporting on reports means that our mechanisms explicitly involve sequential communication, so we refer to them as *sequential communication mechanisms*. It might also be noted at this point that the mechanisms we allow are non-random. Every array of messages leads to a pure action. We will induce randomization when we allow agents to randomize over the messages they send in $[0, 1]$.

An equilibrium for the competing mechanism game is a Bayesian equilibrium of the usual sort. The players' strategies specify for each of their types, a mechanism and a rule that specifies the messages they send in each round as a function of the mechanisms offered by the other players, and the messages they received in previous rounds. When we need it, we use the notation Σ_i to refer to the set of strategy rules available to player i. The notation σ_i refers to a specific strategy. A Bayesian equilibrium is a collection of strategy rules that are jointly best replies to one another.

3. Theorem

At this point we can state our main theorem:

Theorem 1. *If there are* 4 *or more players, then an allocation rule can be supported as a Bayesian equilibrium in the competing mechanism game if it is incentive compatible and individually rational.*

It is important to point out what this theorem adds to the logic in Yamashita. Most obviously it covers random, and even correlated outcomes that could not be captured because of the pure strategy non-random mechanism assumptions in Yamashita. Secondly, it extends the characterization from the uninformed principal informed agents framework to an environment in which there are informed principals. It covers common agency provided there are three or more principals, which is ruled out by Yamashita's approach. It also admits problems in which bargaining power is evenly distributed among players. At the most fundamental level, it provides a characterization in the form of a set of inequalities, which Yamashita's paper does not do, as we explained above.

Of course, Yamashita's point is not to provide a characterization in the first place. It is simply to show how recommendation mechanisms work. Our model goes beyond this. We start with the set of incentive compatible individually rational allocation rules, then show how to implement all of them.[1]

We now turn to the proof of this result which is completely constructive.

4. Some Preliminary Ideas.

Our proof combines a number of ideas. We borrow methods from computer science to implement correlated and random outcomes. We then develop a sequential communication mechanism that effectively converts private communication into a public correlating device. We explain each of these methods before we proceed to the proof of the main theorem.

4.1. **Implementing random outcomes with non-random contracts.** Let B be a set with K elements indexed in some arbitrary way. Let π be a vector of K probabilities that sum to one. Let $\tilde{t}$ be a random variable uniformly distributed on $[0,1]$. The *randomizing function* $\alpha^{\pi}(\cdot, B)$ for mixture π on the set on set B is defined by

$$\alpha^{\pi}\left(\tilde{t}, B\right) = \left\{ b_k : k = \min_{k \in \{1,\ldots,K\}} \sum_{l=1}^{k} \pi_l \geq \tilde{t} \right\}. \tag{4.1}$$

This randomizing function takes value b_k with probability π_k. To see how this device will be used, suppose that player i can observe a verifiable random device $\tilde{t}$ which is uniformly distributed on $[0,1]$. Then the contract $\alpha^{\pi}\left(\tilde{t}, A_i\right)$ which maps from the randomizing device into pure actions implements the mixture π on A_i. More broadly, $\alpha^{\pi}\left(\tilde{t}, A\right)$ implements joint action a^k with probability π_k. Let $\alpha_i^{\pi}\left(\tilde{t}, A\right)$ be the projection of α onto A_i. If each player writes a contract based on $\tilde{t}$ that commits them to take action $\alpha_i^{\pi}\left(\tilde{t}, A\right)$, then the set of contracts $\left\{\alpha_1^{\pi}\left(\tilde{t}, A\right), \ldots, \alpha_n^{\pi}\left(\tilde{t}, A\right)\right\}$ implements the joint randomization π.

4.2. **A property of uniform distributions.** For any non-negative real number x, $\lfloor x \rfloor$ means the *fractional part* of x (sometimes the terminology is x mod 1). Let $\tilde{x}_1, \ldots, \tilde{x}_n$ be a collection of n independent random variables, where each $\tilde{x}_i$ is uniformly distributed on $[0,1]$. For $n \geq 2$, fix $\tilde{x}_i = \overline{x}$ for some i. Then $\lfloor \overline{x} + \sum_{j \neq i} \tilde{x}_j \rfloor$ is a random variable. This random variable turns out to be uniformly distributed on $[0,1]$

[1]Even if we don't know what these allocation rules are, it seems a far easier problem to calculate them for some environment than it does to find all mechanisms which have pure strategy continuation equilibrium.

independent of $\bar{x}$. [2] Since this argument proves very useful below, we give a simple proof in the Appendix section 8.1..

4.3. **Confirmation Process.** Now we describe how we will turn private messages into public messages. There are two issues here - the first is to create what amounts to a public correlating device. Perhaps as important, each player will convey type information to the other players. Since the player's type report must be the same in each mechanism to which he reports, we have to provide some kind of incentive for players to say the same thing to all players. We do this using a special sequential communication mechanism that we call a 'confirmation process'. Players send messages in the first round, but commit themselves to react only when other players confirm these messages in the second round.

Recall that in our mechanisms, a player expects to receive a message in $(\Theta \times [0,1])^n$ from each of the other players along with some recommendation about how to deal with these messages. The recommendation is a mapping $\gamma_i : (\Theta \times [0,1])^{n(n-1)} \to A_i$ suggesting to player i how he should convert the messages in $(\Theta \times [0,1])$ into actions. As they aren't mechanisms per se, we will refer to the mappings γ_i as *processes*. To simplify notation a bit, define $S \equiv (\Theta \times [0,1])$. Let τ be a mapping from $S^{n(n-1)} \to S^n$.

Now adopt the following notation: let s_j represent the first round message that player i receives from player j. Let t_j^k represent the second round report that player j makes to i about player k's first round report. Observe that the message t_j^i would be j's report about what i told him in the first round. This message is important in what follows despite the fact that i already knows what report he made to j in the first round. Player i's mechanism will deal with $n-1$ reports like this and convert them into a single vector in S^n.

Definition 2. The mapping $\gamma_i : S^{n(n-1)} \to A_i$ is called a *confirmation process* for player i if there is a mapping $\tau : S^{n(n-1)} \to S^n$ and a mapping $\gamma_i^\tau : S^n \to A_i$ such that $\gamma_i(\tilde{s}) = \gamma_i^\tau(\tau(\tilde{s}))$ for every $\tilde{s} \in S^{n(n-1)}$, and such that the j^{th} component of this transformation τ is given by

$$\tau_j(s_{-i}, t_{-i}) =$$

[2] This appears to be conventional wisdom in statistics. The theorem is referred to in Deng and E.Olusegun (1990). A proof that the sum mod 1 of a pair of random variables on $[0,1]$ is uniform as long as at least one of the random variables is uniform is given in Deng, Lin, Wang, and Yuan (1997), Theorem 3.1 (see especially the comment after the theorem).

$$
(4.2)\quad \begin{cases} s' & \text{if } j = i; \left\{\exists! j' : t^i_{j'} \neq t^i_k \equiv s' \forall k \neq j', i\right\} \vee \left\{t^i_k = s' \forall k \neq i\right\} \\ s' & \text{if } j \neq i; \left\{t^j_k = s' \forall k \neq j\right\} \vee \left\{\exists! k \neq i, j : t^j_k \neq s_j = s'\right\} \\ \underline{s} & \text{otherwise.} \end{cases}
$$

In the expression above, the notation $\exists!$ means there 'exists a unique...', the notation $\vee$ stands for 'or'.

We explain. A confirmation process breaks the mapping γ_i into two parts. One part, the γ_i^τ looks much like an inscrutable direct mechanism in which the action the player takes is a function of the types of his agents as well as his own type. His 'agents' in this case, are just the other players. However, he doesn't simply ask the others to report their types or choose his own type. Instead, he derives these types from a larger set of messages. This is the 'confirmation' part τ, which reduces the $n\,(n-1)$ messages sent over two rounds into the n messages required by γ_i^τ.

The number $\tau_j\,(s_{-i}, t_{-i})$ is the type (and correlating message) in S that i will use for player j in the 'direct mechanism' γ_i^τ. First consider how player i derives the type that he uses for himself - i.e., the value τ_i. The logic is described in the first line of (4.2). He asks the others to tell him what type he reported to them in the first round. If they all agree, or all but one of them agrees he uses whatever type they agree on. Otherwise, he uses some arbitrary type.

For the others, the computation differs only slightly - the logic is described in the second line of (4.2). To find a type (and correlating message) for player j, he asks the players other than j what type j reported to them in the first round. If they all agree, he uses that type. If there is a single dissenting message, he compares the messages that do agree with the type that j reported to him on the first round. If those agree, then he uses that type. Otherwise, he uses an arbitrary type.

We give more structure to the processes γ_i below. For the moment, we focus on a very specific property of confirmation processes. Fix an array of mechanisms for the players. Every such array of mechanisms indexes a subgame of the original game in which players send recommendations and reports. Generally players don't know what recommendations other players make to each other, so they can't predict exactly how their reports in S^n are being converted into actions by any other player.

However, if we fix a set of strategy rules, then each player j believes that the relationship between reports and actions for player i is given by some mapping $\tilde{\gamma}_i^j : S^{n(n-1)} \to P_i$. This mapping depends on j's type, though we suppress this in the notation to make it a bit simpler.

Implicit in this mapping is a presumption that j follows the strategy σ_j when he makes his recommendation to i.

In many subgames, this uncertainty will disappear. For example, if player i uses a mechanism which makes his action independent of players' recommendations. The case we are interested in here is one in which the strategy rules that players are using are such that player j knows what recommendations the others will make to player i. Given some subgame and some array of strategy rules, we say that player j believes that player i is using process γ_j if $\tilde{\gamma}_j^i = \gamma_j$.[3]

Lemma 3. *Suppose $n \geq 4$. Consider any subgame and set of strategy rules such that some player j believes that player i is using a confirmation process. Suppose further that all the players other than j are using strategy rules that involve a consistent revelation strategy. Then whatever the realizations (s_{-j}, t_{-j}) of the others' reports, $\tau_k^i(s_{-i}, t_{-i})$ is independent of what j reports if $k \neq j$, while there are reports that j can send to i such that $\tau_j^i(s_{-j}, t_{-j})$ takes any value in S.*

The important thing about this Lemma is that when j considers what messages to send to player i, he is strategically in exactly the same position he would be in participating in a standard direct mechanism. All he can affect is his type report (and correlating message) to player i.

We give a full proof here, but the logic is straightforward. There are 4 or more players, so i expects messages from at least three players. By (4.2) i will ignore a message from one player unless it agrees with all the others. When player j is considering what value player i will use for the type and correlating message of some player k, he expects the other players to report a common value in S to i as they are all using consistent reporting strategies. As a consequence, he expects his own message will be ignored. Observe that this logic applies to both the first and second round message. Also note that player i is also expected to use a consistent reporting strategy here, so k could have the value i in this paragraph.

On the other hand, if j considers what message i will use for him, the logic is different. As the others are using consistent reporting strategies, he simply needs to send the same message in S to each of the other players to ensure that i uses that message. None of this argument works if n is three or less because i cannot tell which of two different messages he should ignore.

[3] Equality here means that the mixture on the left is a degenerate mixture with support that coincides with the values on the right hand side.

4.4. **Consensus Mechanisms.** The idea that principals should ask their agents for recommendations about how to process information is due to Yamashita (2010). His idea was to have the principal commit himself to carry out the recommendations of the agents provided an outright majority of the agents make the same recommendation. We simply adapt this idea here. In our context there may or may not be agents, so players ask other players for recommendations. If the principal's mechanism commits him to carry out the recommendation when all the other players, or all but one of the other players agree, then we say that the principal's mechanisms is a *consensus mechanism.*

Formally, a mechanism $r_i : (\Gamma_i)^{(n-1)} \times S^{n(n-1)} \to A_i$ is a consensus mechanism if

$$r_i\left(\gamma_{-i}, s_{-i}\right) =$$

$$\begin{cases} \gamma'\left(s_{-i}\right) & \text{if } \{\exists! j : \gamma_j \neq \gamma_k \equiv \gamma' \forall k \neq j\} \vee \{\gamma_k = \gamma_j \equiv \gamma' \forall j, k\} \\ \overline{a}_i & \text{otherwise.} \end{cases}$$

Now the proof of our folk theorem can be done constructively. On our equilibrium path all players will offer a consensus mechanism independent of their type. If all players do this, then each of them will recommend a confirmation process to each of the other players. The details of the confirmation process will depend on the allocation rule we are trying to support. If some player deviates and offers something other than a consensus mechanism, then the other players will recommend to each other a confirmation process than penalizes the deviator.

It should be apparent why this construction will work. Players can see whether or not everyone has offered a confirmation process after mechanisms are announced. They then believe that they know what recommendations the others will make. The nature of a consensus process is such that unilateral disagreement is ignored, so going along with the majority is at least a weak best reply. The confirmation process has been constructed specifically so that all players can accomplish when they participate is to report one common type to all the others. As long as the confirmation process implements an incentive compatible allocation rule, it is a best reply for each player to report this type truthfully. The only real complication in the proof is to show that it is an equilibrium for players to send correlating messages that correctly implement randomized outcomes.

5. The Proof of the Main Theorem

Proof. The proof is constructive. Let $q(\theta)$ be the randomization that is to be supported when types are θ. Since the allocation rule is individually rational, there is a collection of punishments that ensure participation by each player. Let $\{p_i(\theta_{-i})\}_{i \in N}$ be the type contingent randomization that is to be carried out by the players other than i when i is being punished.

We first describe the recommendations we want players to make.

Let τ be a confirmation process with message space $S^n = (\Theta \times [0,1])^n$. Write (θ, x) as a typical element of S. The function $\tau_j(s_{-i}, t_{-i}) \in \Theta \times [0,1]$, so write $\tau_j(s_{-i}, t_{-i}) = \left\{\tau_j^\theta(s_{-i}, t), \tau_j^x(s, t_{-i})\right\}$. The equilibrium path recommendation by other players to player i is given by

$$\gamma_i(s_{-i}, t_{-i}) = \alpha_i^{q\left(\tau^\theta(s_{-i}, t_{-i})\right)}\left(\lfloor \sum_{j \in N} \tau_j^x(s_{-i}, t_{-i}) \rfloor, A\right) \tag{5.1}$$

where α_i^q is the projection of the randomizing function for mixture $q\left(\tau^\theta(s_{-i}, t_{-i})\right)$ on the set A onto the set A_i. The randomizing function is defined by (4.1) above.

When player k unilaterally deviates in the mechanism design stage and offers something other than a recommendation mechanism, the non-deviators will recommend

$$\gamma_i^k(s_{-ik}, t_{-ik}) = \alpha_i^{p_i\left(\tau^\theta(s_{-ik}, t_{-ik})\right)}\left(\lfloor \sum_{j \neq k} \tau_j^x(s_{-ik}, t_{-ik}) \rfloor, A_{-k}\right) \tag{5.2}$$

to each non-deviating player i, where (s_{-ik}, t_{-ik}) is an array of messages from the other *non-deviating* players. In words, the non-deviators will recommend to each other the (projection of the) randomizing function associated with the punishment.

In any information set in which all players have offered a consensus mechanism, player i should recommend γ_j to each other player j, truthfully report to each player $k \neq j$ the message received from player j, truthfully report his type to every other player, and send every other player a correlating message x drawn uniformly from $[0,1]$. In any history in which a single player, say player k, has deviated and offered some mechanism other than a consensus mechanism, player i should recommend the punishment mechanism γ_j^k to each player $j \neq k$, truthfully report the private message received from each player $j \neq k$ to each player $j' \neq k, j$, send the same correlating message s' to each of the other players where s' is chosen using a uniform distribution on $[0,1]$, and report his type truthfully to each of the players other than k. Any action unspecified here can be chosen arbitrarily.

Now we proceed to prove that the strategies specified constitute a Bayesian equilibrium. First, it is immediately a best reply for each player to offer a consensus mechanism. If he does that, he should expect the allocation rule $q(\theta)$. If he deviates, he should expect the others to implement the punishment $p_i(\theta)$. Since the allocation rule satisfies (1.2), this can't increase his payoff.[4]

On the equilibrium path, all players offer consensus mechanisms, and each player recommends a confirmation processγ_j as given by (5.1). We have already explained that for any confirmation process, it is a best reply for each player to report the same type and correlating message to each of the other players provided they believe that the others are doing the same. For the correlating message, the others are expected to send a correlating message that is uniformly distributed on $[0,1]$. As we have explained in Remark 4 above, this implies that $\lfloor \sum_{j\in N} \tau_j^x \rfloor$ has a uniform distribution independent of what signal x_i i chooses to send. Then for each θ_{-i} and each report θ_i' and signal $\tilde{x}_i$ that i chooses to send to the others on the first round

$$\alpha_k^{q\left(\tau^\theta\right)}\left(\lfloor \sum_{j\in N} \tau_j^x \rfloor, A\right) =$$

$$\alpha_k^{q\left(\theta_i',\theta_{-i}\right)}\left(\tilde{x}, A\right)$$

where $\tilde{x}$ has a uniform distribution on $[0,1]$. Since this rule implements the incentive compatible rule q, player i has no incentive to misrepresent his type. It is also a best reply for player i to choose a signal uniformly from $[0,1]$. $\square$

6. Remarks.

The approach above shares many of the methods of the literature on communication in gamesGerardi (2004),Forges (1986),Barany (1992). Gerardi (2004), for example, uses the majority rule approach to ensure that players all send the 'correct' message in his communication protocols. This is exactly the idea behind a consensus mechanism. He also uses the randomization idea in (4.1), albeit restricted to two players.[5] The important difference between our paper and all this literature is

[4]Notice that because we are only interested in Bayesian equilibrium at this point, this particular argument works even if there are only three players in the game. We still need the fourth player to support equilibrium in the confirmation process.

[5]He has two players publicly announce numbers in the interval $[0,1]$ then uses the fractional part as a public correlating device. As there are only two players,

the fact that we are doing mechanism design - players can make commitments based on messages. So the allocation rules we support aren't typically communication equilibrium (or correlated equilibrium with complete information).

As an example, consider a prisoner's dilemma played between two players 1 and 2. To make the environment fit our settings, add two disinterested players 3 and 4 who take no actions of their own. The actions are C for cooperate and D for defect. The only communication equilibrium in this game has both players A and B playing D, since the action C is striclty dominated. The outcome where both 1 and 2 play C can be supported as an equilibrium with recommendation mechanisms. A recommendation mechanism commits the player to the action the other three players recommend provided 2 of the three recommendations agree. To keep things simple, suppose the only other mechanisms that players are allowed to offer are the ones that ignore all messages and commit to either C or D. The strategies are for each player 1 and 2 to offer a recommendation mechanism then recommend C if the other player offers a recommendation mechanism, Players 3 and 4 recommend C if 1 and 2 both offer recommendation mechanisms, and recommend D if one player offers a recommendation mechanism and the other doesn't. It should be apparent in this construction that deviating from this equilibrium changes the action of the other player from C to D. So these strategies constitute an equilibrium.

What is important in this exercise is that players 1 and 2 have a way to *commit* themselves to an action which can never be part of a communication equilibium.

There is a literature on mechanism design in communication networks (J. Renault and Tomala (2010) or Renou and Tomala (2010)) which considers sequential communication schemes like the one we described in Section 4.3. In this literature, a centralized mechanism designer can communicate with only a subset of all the agents. However, the agents can communicate among themselves according to some exogenously fixed communication protocol. The papers cited above provide communication protocols which allow agents to communicate their type information secretly to the principal. The essence of their result is to show that, provided the communications network is right, there is a way for agents to encode their own information along with the information they have received from others, and pass it along in such a way that only the the mechanism designer can decode it.

he doesn't need our Remark 4. In our model, there are no public messages at all, beyond the mechanisms that players announce at the beginning of the game.

In order to ensure that players pass along encoded information truthfully, their papers use a method that resembles our confirmation process. A protocol that transmits player 1's type (assumed here to be a positive number) to the mechanism designer is repeated, say, 3 times. Player 1 chooses at random one of the three repetitions and transmits his type on that repetition. On each of the other two repetitions he transmits the number 0 as his type. If the mechanism designer decodes 2 zeros and one positive number, he responds as if the type is a positive number. If he decodes any other sequence, he implements a punishment. The purpose of this is to ensure that the other players transmit messages from player 1 truthfully. They don't know which of the three messages from player 1 contain his type report. So they have a 1 in three chance of changing the outcome in a way that they might like, and a 2/3 chance of inducing the punishment when they lie. Assuming that there is a punishment that is strictly worse than any outcome the mechanism designer might otherwise implement, no matter the type of any other player, then repeating the protocol enough times will ensure that players other than player 1 transmit messages truthfully.

The details of the argument differ, but the spirit is the same as the confirmation process - if other players are hearing the same message that you are, then there are sometimes ways to to check whether they are transmitting the information truthfully. Our communication process is structured to do this, so we don't need a 'worst outcome' that a mechanism designer can use to enforce truthtelling. In our framework, deviating messages are simply ignored. Of course, the context of our result is quite different since we don't have a mechanism designer in the first place - we deal with decentralized competition.

Nonetheless, the method they describe illustrates how the results presented here might be extended to games with fewer than four players. Our communication mechanism requires agreement among all but one of the players who are participating in a mechanism. If there are only two players, and the messages they send are different, then the player who is interpreting them does not know which message is the correct one, and which is a deviation. Each of our players has to have at least three others sending him messages for our method to work. The method above illustrates how a player might detect deviations with messages from only two players provided the sequential communication mechanism goes on for long enough. It may also be possible to extend our results if there is a public correlating device using methods like those in Forges and Vida (2011) who show that communications equilibrium outcomes can be implemented with long cheap talk in games with only two players using a public device.

The use of a second round of communication to provide a mechanism designer with additional information is similar to the argument in Mezzetti (2004), who shows how a mechanism designer can improve outcomes by using a second round of information in which players provide information about their values. When players' payoffs are interdependent, each player's value contains information about everyone else's type in much the same way the first round reports do here. Of course, the method we use to get players to reveal this information is quite different than it is in that reference.

Folk theorem like results for competing mechanism games have been provided by Tennenholtz (2004), Kalai, Kalai, Lehrer, and Samet (2010) and Peters and Szentes (2012). The essential difference between these papers and our result here is that they assume contracts condition directly on the contracts of other players. The paper by Peters and Szentes (2012) deals with incomplete information games. It fully characterizes the outcome functions that can be supported as contract equilibrium. However, it assumes that players never communicate privately. Any type information that a player wants to convey must be publicly conveyed through his contract offer. This can limit the effectiveness of punishments since a deviating player will inevitably know the types of the other players when he deviates. It is difficult to give a formal description of the difference between the two papers because Peters and Szentes (2012) rule out randomization. To illustrate the relationship between the outcome function and the information that a deviator would then have during the punishment phase, we would need to develope considerable additional formalism. Roughly speaking, their characterization provides an individual rationality constraint that looks like 1.2 except for the fact that the deviator's beliefs when he chooses his best action would depend on the types of the punishing players. They provide an example of an outcome function that is supportable in the sense described here, which cannot be supported as an equilibrium in their game because of the fact that firms equilibrium contract offers leak information about their types. So the set of outcome functions supportable as Bayesian equilibrium in the Peters and Szentes (2012) model is strictly smaller than the set supported here.

The paper by Peters (2010) provides a characterization of outcome functions supportable as Bayesian equilibrium in regular contracting games. Along with the characterization it provides a contracting game in which these outcomes can be supported as *perfect* Bayesian equilibrium. In fact it borrows the correlating device we have, in turn, borrowed here. However, the point of that paper is quite different. It revisits the question in Epstein and Peters (1999) and provides a modified set of direct mechanisms that can be used to mimic equilibrium outcomes in any competing mechanism game - effectively providing a revelation principle for competing mechanisms.

The set of outcome functions supported here is a subset of the set supported in Peters (2010). The reciprocal contracting game described in Peters (2010) supports more outcomes because it adds a cheap talk stage to the game that isn't present in Yamashita's game. The makes it possible to support some punishments in which the deviating player's action depends on the punishing players' types (which is impossible in the framework here). Enlarging the set of possible punishments enlarges the set of outcome functions that can be supported in equilibrium.

It would be possible to modify Yamashita's game (adding cheap talk and proceeding exactly as in Peters (2010)) so that all the outcomes described in Peters (2010) could also be supported. However, there is little point in doing so. The full set of outcome functions supportable in Yamashita's game is unknown for reasons we explain in the next section. Our theorem above does not provide a full characterization of supportable outcomes.

7. Bayesian Equilibrium

Our theorem uses Bayesian equilibrium as a solution concept. Beyond saying that refined equilibrium will impose additional restrictions, not much can be said about perfect Bayesian equilibrium in Yamashita's game.

To understand the biggest problem with refinements, consider a game with complete information. There are four players (simply so that the assumptions of our theorem above are satisfied). Suppose that player 1 has three possible actions, $\{a, b, c\}$. None of the other players controls any actions at all. Player 1 offers a mechanism, and the solution concept requires that after seeing the mechanism, continuation play constitutes a Nash equilibrium (subgame perfection). Obviously, player 1 simply chooses his favorite action in any Bayesian equilibrium. However, player 1 could deviate and offer a mechanism which invites players 2 and 3 to send a message in $[0, 1]$. He commits to translate the messages m_2 and m_3 into actions the following way:

$$\gamma(m_2, m_3) = \begin{cases} a & \text{if } m_2 < m_3 < m_2 + \frac{1}{2}, \\ b & \text{if } m_2 = m_3 \text{ or } m_3 = m_2 + \frac{1}{2}, \\ c & \text{otherwise.} \end{cases}$$

Now imagine payoffs for player 2 are$u(a) = -1$, $u(b) = 0$, and $u(c) = 1$. Player 3's payoff is $-u$. This is simply the Sion Wolfe Sion and Wolfe (1957) example of a game that has no equilibrium in either pure or mixed strategies. This is a feasible mechanism in our framework, and a reasonable looking mechanism in any framework. So in this simple setting, there can be no subgame perfect equilibrium to the mechanism game unless mechanisms like the one above are ruled out.

One approach is to restrict the set of mechanisms that players are allowed to recommend to principals (by requiring that mechanisms only use finite message spaces for example so that continuation equilibrium always exists). An alternative approach would be to use a refinement other than subgame perfection.[6] For example, Peters and Szentes (2012) suggest a refinement that looks more like sequential rationalizability. Either approach would impose additional restrictions on the punishments.

There is little point in trying to develop these approaches, because recommendation games suffer from another difficulty, even when the solution concept is Bayesian equilibrium. The issue is that agents see what contract a deviator has offered before they make a recommendation about how to punish. As we have modelled the game here, a mechanism offered by a deviator must specify an action. As a consequence, Yamashita's game is not *regular* as defined in Peters (2010), since the players who are punishing the deviator can make their actions depend on what action the deviator commits to play. Punishments could then have a kind of maxmin property.

This seems an undesirable property of a competing mechanism game because it is effectively eliminating an option that is available to players in the default game. To put it another way, adding a contracting process that involves recommendation mechanisms amounts to changing the economic environment in which players interact.

This is not such a big problem in games of complete information since there is no difference between minmax and maxmin. However in the incomplete information environments discussed here, it just isn't clear what the implication of this change in the economic environment will be. Indeed, the whole idea of 'punishing' a deviator with an incentive compatible (or incentive rationalizable) punishment breaks down in Yamashita's game. Whether a punishment is incentive compatible or not is going to depend on what the deviator chooses to commit to. A full characterization of equilibrium would then require writing

[6] Sequential equilibrium is not well suited to the game discussed here because the messages spaces aren't finite.

down a list of punishment not only for every deviator, but for every commitment the deviatior might make.

The implication of all this is that there are outcomes supported as Bayesian equilibrium in Yamashita's game, that cannot be supported with reciprocal contracting as in Peters (2010). Indeed there must be outcomes that could not even be supported by a centralized mechanism designer who is constrained to select punishments as they are usually described. It is just very difficult to understand what these outcome functions would look like.

Conclusion

Our basic contribution is to show that a large set of outcomes can be supported as equilibrium in Yamashita's recommendation game. The methods used to support randomization, and to coordinate the information available to the players should be useful in other contexts.

8. Appendix

8.1. **Uniform Distributions and independence.**

Remark 4. $\lfloor \bar{x} + \sum_{j \neq i} \tilde{x}_j \rfloor$ is uniformly distributed on $[0,1]$ independently of $\bar{x}$ provides each $\tilde{x}_j$ is uniformly distributed on $[0,1]$.

Proof. Suppose that $n = 2$. Then $\sum_{j \neq i} \tilde{x}_j = \tilde{x}_j$, and $\lfloor \bar{x} + \tilde{x}_j \rfloor$ is obviously uniform. Let both $\tilde{x}_1$ and $\tilde{x}_2$ be uniform on $[0,1]$. Then the probability density function of $\tilde{z} = \tilde{x}_1 + \tilde{x}_2$ is[7]

$$f(z) = \begin{cases} z & 0 \leq z \leq 1 \\ 2 - x & \text{otherwise.} \end{cases}$$

The probability that $\lfloor \tilde{z} \rfloor \leq w$ is then given by

$$\int_0^w z dz + \int_1^{1+w} (2 - z)\, dz = w.$$

So $\lfloor \tilde{x}_1 + \tilde{x}_2 \rfloor$ is uniformly distributed. So when $n = 3$, $\lfloor \bar{x} + \sum_{j \neq i} \tilde{x}_j \rfloor$ is uniformly distributed. Then the argument follows by induction. If for $n - 1$ players $\lfloor \bar{x} + \sum_{k \neq j} \tilde{x}_k \rfloor$ is uniformly distributed, then for n players

$$\lfloor \bar{x} + \tilde{x}_j + \sum_{k \neq i,j} \tilde{x}_k \rfloor =$$

$$\lfloor \tilde{x}_j + \lfloor \bar{x} + \sum_{k \neq i,j} \tilde{x}_k \rfloor \rfloor$$

[7] Hall (1927).

and uniformity follows from the result for $n = 3$. □

8.2. **Proof of Lemma 3.**

Lemma 5. *Suppose $n \geq 4$. Consider any subgame and set of strategy rules such that some player j believes that player i is using a confirmation process. Suppose further that all the players other than j are using strategy rules that involve a consistent revelation strategy. Then whatever the realizations (s_{-j}, t_{-j}) of the others' reports, $\tau_k^i(s_{-i}, t_{-i})$ is independent of what j reports if $k \neq j$, while there are reports that j can send to i such that $\tau_j^i(s_{-j}, t_{-j})$ takes any value in S.*

Proof. Fix the first round reports s_{-j} of the players other than j. We write in the obvious way s_{-jk} for the subvector consisting of reports in s_{-j} by players other than k. Suppose that j's strategy is consistent and he sends the message s' to each of the other players in the first round. Then since every other player is using a consistent strategy, the value that i uses for player j will be based on first round messages (s', s_{-ij}), second round message (s', s_{-kj}) from each player $k \neq j$ since each such player is using a consistent reporting strategy, and second round message s_{-j} from player j which doesn't depend on s'. Since the first round message from player j agrees with the second round reports of each of the other players, we conclude by (4.2) that

$$\tau_j^i \left((s', s_{-ij}), \prod_{k \neq i,j} (s', s_{-kj}), s_{-j} \right) = s'.$$

Notice that this verifies the last part of the theorem - j can induce any value for τ_j^i in S.

Player j can deviate from this consistent strategy by sending different messages to the other players on the first round. He could also send different messages to i on the second round, but τ_j^i doesn't depend on i's second round messages, so we defer discussion of this second kind of deviation. Let s_k be the message he sends to player k on the first round, and $\tilde{s}_{-j}$ the vector of $n-1$ messages he sends to i on the second round. In this case there are two possibilities. If the players $k \neq j$ all report $s'_k = s'$, or if all but one of the others reports $s' = s'_i$ then by (4.2),

$$\tau_j^i \left((s'_i, s_{-ij}), \prod_{k \neq i,j} (s'_k, s_{-kj}), \tilde{s}_{-j} \right) = s',$$

which is an outcome j could have obtained by using a consistent reporting strategy and reporting s' in the first round to everyone. Otherwise,

$$\tau_j^i \left(\left(s_i', s_{-ij}\right), \prod_{k \neq i,j} \left(s_k', s_{-kj}\right), \tilde{s}_{-j} \right) = \underline{s},$$

which is an outcome he could also accomplish with a consistent strategy by sending the message $\underline{s}$ to each player then reporting accurately to i in the second round.

To complete the proof of the theorem, observe that since player k is using a consistent reporting strategy, he will make the same first round report s_k to each of the other players. With the possible exception of player j, each of the others will then report s_k to player i. Since at least two second round reports will agree with k's first round report, we have

$$\tau_k^i \left(\left(s_i', s_{-ij}\right), \prod_{k' \neq i,j} \left(s_{k'}', s_{-k'j}\right), \tilde{s}_{-j} \right) = s_k$$

independent of $\tilde{s}_{-j}$. □

References

Barany, I. (1992): "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research*, 17, 327–340.

Deng, L.-Y., and G. E.Olusegun (1990): "Generation of Uniform Variates from Several Nearly Uniformly Distributed Variables," *Communications in Statistics-Simulation and Computation*, 19(1).

Deng, L.-Y., D. K. Lin, J. Wang, and Y. Yuan (1997): "Statistical Justification of Combination Generators," *Statistica Sinica*, 7, 993–1003.

Epstein, L., and M. Peters (1999): "A Revelation Principle for Competing Mechanisms," *Journal of Economic Theory*, 88(1), 119–160.

Forges, F. (1986): "An Approach to Communication Equilibria," *Econometrica*, 54, 1375–1385.

Forges, F., and P. Vida (2011): "Implementation of Communication Equilibria by Correlated Cheap Talk: The Two Player case," Discussion paper, CESifo Working Paper No. 3360.

Gerardi, D. (2004): "Unmediated communication in games with complete and incomplete information," *Journal of Economic Theory*, 114(1), 104–131.

Hall, P. (1927): "The Distribution of Means for Samples of Size N drawn from a Population in which the Variate takes Values between 0 and 1, All Such Values Being Equally Probable," *Biometrika*, 19(3), 24–245.

J. Renault, L. R., and T. Tomala (2010): "Secret Information Transmission," Discussion paper, HEC Paris.

Kalai, A., E. Kalai, E. Lehrer, and D. Samet (2010): "A Commitment Folk Theorem," *Games and Economic Behavior*, 69(1), 127–137.

Mezzetti, C. (2004): "Mechanism Design with Interdependent Valuations: Efficiency," *Econometrica*, 72, 1617–1626.

PETERS, M. (2010): "On the Revelation Principle and Reciprocal Mechanisms in Competing Mechanism Games," Discussion paper, University of British Columbia.

PETERS, M., AND B. SZENTES (2012): "Definable and Contractible Contracts," *Econometrica*, 80(1), 363–411.

RENOU, L., AND T. TOMALA (2010): "Mechanism Design and Communication Networks," Discussion paper, HEC Paris no. 926/2010.

SION, M., AND P. WOLFE (1957): *On a Game without a Value*Princeton:Princeton University Press.

TENNENHOLTZ, M. (2004): "Program Equilibrium," *Games and Economic Behavior*, 49(2), 363–373.

YAMASHITA, T. (2010): "Mechanism Games with Multiple Principals and Three or More Agents," *Econometrica*, 78(2), 791–801.